View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Aberystwyth Research Portal

Aberystwyth University

# *Breast Image Pre-processing For Mammographic Tissue Segmentation*

He, Wenda; Hogg, Peter; Juette, Arne; Denton, Erika R. E.; Zwiggelaar, Reyer

*Published in:*
Computers in Biology and Medicine

*DOI:*
10.1016/j.compbiomed.2015.10.002

*Publication date:*
2015

*Citation for published version (APA):*
He, W., Hogg, P., Juette, A., Denton, E. R. E., & Zwiggelaar, R. (2015). Breast Image Pre-processing For Mammographic Tissue Segmentation. Computers in Biology and Medicine, 67, 61-73. https://doi.org/10.1016/j.compbiomed.2015.10.002

**General rights**
Copyright and moral rights for the publications made accessible in the Aberystwyth Research Portal (the Institutional Repository) are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the Aberystwyth Research Portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the Aberystwyth Research Portal

**Take down policy**
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

tel: +44 1970 62 2400
email: is@aber.ac.uk

Download date: 03. Oct. 2019

Author's Accepted Manuscript


www.elsevier.com/locate/cbm

# Breast Image Pre-processing For Mammographic Tissue Segmentation

Wenda He, Peter Hogg, Arne Juette, Erika R.E. Denton, Reyer Zwiggelaar

PII: S0010-4825(15)00336-4
DOI: http://dx.doi.org/10.1016/j.compbiomed.2015.10.002
Reference: CBM2252

To appear in: *Computers in Biology and Medicine*

Received date: 8 May 2015
Revised date: 22 September 2015
Accepted date: 2 October 2015

Cite this article as: Wenda He, Peter Hogg, Arne Juette, Erika R.E. Denton and Reyer Zwiggelaar, Breast Image Pre-processing For Mammographic Tissue Segmentation, *Computers in Biology and Medicine* http://dx.doi.org/10.1016/j.compbiomed.2015.10.002

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

# Breast Image Pre-processing For Mammographic Tissue Segmentation

Wenda He[a,1,*], Peter Hogg[b], Arne Juette[c], Erika R. E. Denton[c], Reyer Zwiggelaar[a,**]

[a]*Department of Computer Science, Aberystwyth University, Aberystwyth, SY23 3DB, UK*
[b]*School of Health Sciences, University of Salford, Salford, M6 6PU, UK*
[c]*Department of Radiology, Norfolk & Norwich University Hospital, Norwich, NR4 7UY, UK*

**Abstract**

During mammographic image acquisition, a compression paddle is used to even the breast thickness in order to obtain optimal image quality. Clinical observation has indicated that some mammograms may exhibit abrupt intensity change and low visibility of tissue structures in the breast peripheral areas. Such appearance discrepancies can affect image interpretation and may not be desirable for computer aided mammography, leading to incorrect diagnosis and/or detection which can have a negative impact on sensitivity and specificity of screening mammography. This paper describes a novel mammographic image pre-processing method to improve image quality for analysis. An image selection process is incorporated to better target problematic images. The processed images show improved mammographic appearances not only in the breast periphery but also across the mammograms. Mammographic segmentation and risk/density classification were performed to facilitate a quantitative and qualitative evaluation. When using the processed images, the results indicated more anatomically correct segmentation in tissue specific areas, and subsequently better classification accuracies were achieved. Visual assessments were conducted in a clinical environment to determine the quality of the processed images and the resultant segmentation. The developed method has shown promising results. It is expected to be useful in early breast cancer detection, risk-stratified screening, and aiding radiologists in the process of decision making prior to surgery and/or treatment.

*Keywords:* mammographic segmentation, risk assessment, density classification, peripheral enhancement, BI-RADS, Tabár.

Accepted manuscript

## 1. Introduction

Breast cancer is the most frequently diagnosed cancer in women [1]. To date, the most effective way to overcome the disease is through early detection, precise identification of women at risk, and application of appropriate disease prevention measures [2]. Mammography is the gold standard method in detection of early stage breast cancer before abnormalities become clinically palpable. Within screening mammography, full field digital mammography (FFDM) has become more popular and is gradually replacing screen film mammography (SFM). Many digital mammography units produce images in two forms; 'raw' and 'processed' images. Raw data is often not archived in clinical practice, whilst the appearances of processed (for presentation) images may vary due to different post-processing algorithms applied by mammography manufacturers. A significant amount of dynamic range provided by FFDM systems is redundant after these logarithmic based post-processing. This may result in lower visibility of breast parenchyma in peripheral areas, and large intensity discrepancy between thicker tissue near the chest wall and peripheral areas; see examples in Figure 1.

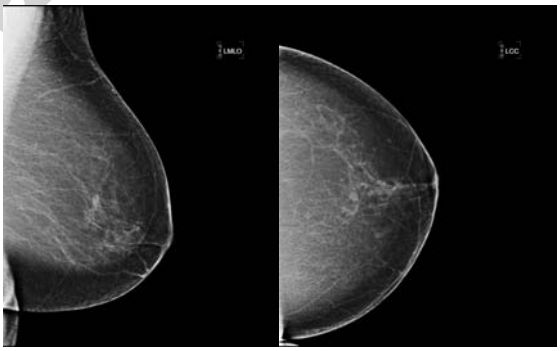


Figure 1: A compression paddle is used to even out the breast tissue during imaging, however, the peripheral areas may not be fully compressed due to a reduction of breast thickness. This results in air gaps above and beneath the uncompressed areas, leading to a non-uniform exposure and degradation in contrast in these areas.

With such processed images, abnormalities near peripheral areas with less visible structures may be missed during a mammogram examination. When used in computer aided mammography, processed images can lead to less satisfactory breast tissue segmentation, due to inter-fatty/dense tissue intensity variation across the mammograms which jeopardises subsequent analysis in the workflow.

A mammographic pre-processing technique can be developed to enhance the visibility of peripheral areas and improve intensity distribution, in order to facilitate interpretation and

*Corresponding author
**Principal corresponding author
*Email addresses:* weh@aber.ac.uk (Wenda He), p.hogg@salford.ac.uk (Peter Hogg), arne.juette@nnuh.nhs.uk (Arne Juette), erika.denton@nnuh.nhs.uk (Erika R. E. Denton), rrz@aber.ac.uk (Reyer Zwiggelaar)

ACCEPTED MANUSCRIPT

benefit follow up analysis. Existing methods in the literature can be categorised into two groups; non-parametric (*e.g.* [3, 4, 5, 6]) and parametric (*e.g.* [7] and [8]) approaches. Most existing methods are intended to be used on 2D mammographic projections. As technology advances, more breast thickness equalisation/correction methods have emerged for 3D volumetric breast density analysis (*e.g.* [9, 10, 11]). The proposed approach is in the application domain of 2D mammograms.

An early non-parametric method [3] focused on balancing the mammographic intensity between the breast centre and its peripheral areas, so that the two areas have 'matching' average greylevel values. A log-like-intensity characteristic curve is created based on the average greylevel values that are within the same distances to the skinline, from which a reversal fitted enhancement curve is obtained using a polynomial fit. This fitted enhancement curve defines the necessary correction value for each pixel, which is added to the original pixel value to create the intensity balanced ('equalised') image. Such an approach applies the thickness correction to the entire breast. It works well with a homogeneous fatty/dense breast but localised artefacts can be seen when a breast exhibits large density variation across the mammogram. To better identify breast peripheral areas requiring correction, a large Gaussian filter can be used to blur a mammographic image isotropically first [4], before obtaining a representation of tissue thickness differences with smoother variations, assuming that the breast thickness variations are smoother than tissue density variations. The thickness correction is only applied in the breast periphery determined by a local threshold at the boundary of the compressed and uncompressed part of the breast. To ensure intensity continuity, a locally determined correction factor is used to multiply with the original pixel values, to derive the corrected pixel values in the breast periphery. It is an effective method to correct pixels in the peripheral areas using neighbouring pixels, however, the corrections can be over emphasised with breasts, which have intricate parencymal structures in the periphery areas, and less desirable corrections can be associated with ringing artefacts. To better reflect breast thickness differences, a mammogram is iteratively segmented into fatty and dense areas prior to the correction, and followed by a linear interpolation to replace all the dense tissue with nearby fatty tissue [5]. An alternative anisotropic diffusion filter based approach (direction parallel to the skin edge) was investigated to facilitate the breast thickness estimation. The method processes the entire breast but only adds correction terms to the pixel values in the peripheral areas. Results showed improved peripheral texture appearances for those structural texture with orientations (*e.g.* blood vessels), however, the interior part of the breast may display higher contrast after correction. The method critically depends on accurate iterative segmentation of the dense breast tissue, which can be problematic when the breast exhibits heterogeneous dense tissue. The aforementioned studies used pixel intensity values as correlation factors to estimate the breast thickness, which may not be an accurate estimation/close approximation. Note that true breast thickness may not be attainable retrospectively.

A parametric method as proposed in [7] used a geometric model of the three-dimensional shape of the breast. The breast interior is modelled by two non-parametric planes which requires three degrees of freedom; one for the onset and two for the slopes. The peripheral area is modelled by bands of semicircles, determined by the breast outline and interior model. Once the parameters of the geometric model are obtained, dense tissue is separated and interpolated with fatty tissue, similar to [5]. Therefore, the breast can be modelled with the original and interpolated fatty tissue. The subsequent correction process is performed by adding a fatty tissue component in the periphery which fills in the air gap between the fitted planes and breast. As in [5], the approach is critically dependent on the accuracy of iterative dense breast tissue segmentation and fatty tissue interpolation. Note that the approach is designed for unprocessed digital mammograms with a linear relationship between exposure and greylevel value, therefore, it cannot be applied to processed FFDM nor SFM with unknown calibration data, which has a non-linear relationship between exposure and greylevel value.

We propose a mammographic pre-processing technique which has the following key novelty aspects: 1) modelling a breast thickness based on its shape outline derived from Mediolateral Oblique (MLO) and Cranio-Caudal (CC) views, instead of using an assumed correlation between smoothed pixels and breast thickness; 2) using a selective approach to target specific mammograms more accurately; and 3) both breast interior and exterior are enhanced simultaneously, in order to achieve intensity balancing across the mammogram and increasing breast tissue visibility in the peripheral areas. Mammographic segmentation and risk/density classification were conducted to determine the usefulness of the developed approach and the results were evaluated in a clinical environment.

With respect to mammoraphic risk assessment, Tabár *et al.*[12] proposed a model based on a mixture of four mammographic building blocks representing the normal breast anatomy, five mammographic risk categorises were identified based on these building blocks (i.e. [nodular%, linear%, homogeneous%, radiolucent%]); $T_I$ [25%, 15%, 35%, 25%], $T_{II/III}$ [2%, 14%, 2%, 82%], $T_{IV}$ [49%, 19%, 15%, 17%], and $T_V$ [2%, 2%, 89%, 7%] [12]. Alternatively, the American College of Radiology's Breast Imaging Reporting and Data System (BI-RADS) [13] was developed, with four breast dense tissue compositions categorised as; $B_1$ the breast is almost entirely fat ($< 25\%$ glandular), $B_2$ the breast has scattered fibroglandular densities ($25\% - 50\%$), $B_3$ the breast consists of heterogeneously dense breast tissue ($51\% - 75\%$), and $B_4$ the breast is extremely dense ($> 75\%$ glandular).

## 2. Data and Method

A Hologic Selenia Dimensions 2D FFDM system was used to obtain a total of 360 digital mammograms (i.e. 180 CC and 180 MLO views), processed for optimal visual appearance for radiologists. Two consultant radiologists[1] provided Tabár risk

[1]One radiologist has over 5 years mammographic reading experience, the other has over 10 years mammographic reading experience.



Accepted manuscript

ACCEPTED MANUSCRIPT

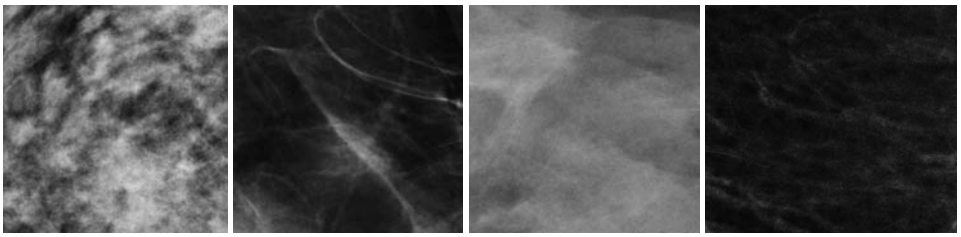


Figure 2: From left to right, showing mammographic patches containing tissue examples for nodular, linear, homogeneous, and radiolucent tissue.

classifications and BI-RADS density ratings for all the mammograms as 'ground truth'. To model mammographic building blocks (breast tissue), a collection of patches were cropped from randomly selected images from the dataset, consisting of examples of (139) nodular, (224) linear structure, (87) homogeneous, and (89) radiolucent tissue; see Figure 2 for examples.

An overall workflow of the developed approach can be found in Figure 3.

### 2.1. Automatic Image Selection

Image acquisition parameters used during screening have direct influences on mammographic appearances. Many factors may cause contrast degradation in breast peripheral areas; such as organ dose, entrance dose, exposure, relative X-ray exposure, compression force, body part thickness, and kVp (peak kilovoltage). Empirical observations showed that no single parameter can be used to determine which image requires pre-processing, although some parameter combinations may provide a better indication than others. Machine learning techniques were employed to build a probability model based on the calibration parameters to determine which mammographic images are more problematic than others. In addition to the calibration parameters obtained from DICOM headers, three image based attributes were calculated from breast peripheral areas, which were segmented automatically using the Otsu algorithm [14]; see Section 2.2.1 for details. The attributes were incorporated as part of multiple variables for each image, including percent peripheral area (PPA), percent pectoral coverage (PPC) of the image rows, and percent skinline coverage (PSC) of the image rows. When calculating the PPC and PSC row coverage within the breast area, a row is valid and counted if a least 15% of the pixels in that row are segmented by the Otsu algorithm. The threshold value was empirically defined based on its ability to correctly determine an image requiring pre-processing; see Figure 4 for examples. Therefore, for each image, a total of ten attributes were collected as features for analysis. Note that extensive PPC or a small amount of PSC is an indication of less abrupt intensity changes and the original image already has a balanced intensity distribution and may not need peripheral correction. The probability model was trained in Weka [15] using 72 images. Note that the training data has two subsets; 50% were randomly selected from images requiring pre-processing, and the other 50% were randomly selected from 'normal' (not problematic) images. The distributions for the two subsets of images labelled with $B_1$ to $B_4$ are in line with the entire dataset. All available classifiers (see full list in [15]) were evaluated and validated by averaging over 10 repeats. A classifier built using the random forest algorithm performed the best and was chosen for the experiment. Table 1 shows statistics for the calibration parameters and the three image based parameters based on the entire dataset.

### 2.2. Pre-processing

Image pre-processing consists of four stages; 1) breast periphery separation, 2) intensity ratio propagation, 3) breast thickness estimation, and 4) intensity balancing.

#### 2.2.1. Breast Periphery Separation

To separate the breast peripheral area ($BPA$), Otsu thresholding [14] is employed and only applied to the breast region. Breast masks were provided as part of the dataset; there are well established automatic methods to extract breast regions which can be found in the literature. The correctness of the initial segmentation ($BPA_{Otsu}$) may not be accurate as the algorithm can miss certain peripheral areas (i.e. the bottom half area in Figure 6 (b)). To improve the binary segmentation, the original image ($Img$) is further thresholded ($BPA_{threshold}$). The threshold value ($T$) is determined as the mean pixel greylevel value for $BPA_{Otsu}$, where $BPA_{threshold} = Img \leq T$. Therefore, the final $BPA = BPA_{Otsu} \cup BPA_{threshold}$, *e.g.* Figure 6 (c). Morphological filling [2] is applied to $BPA$ to fill small holes, followed by morphological dilation to connect close neighbouring pixels. Note that both morphological operations used a small $3 \times 3$ structuring element. The $BPA$ is refined by only keeping the single biggest connected component; this is done by iteratively labelling connected neighbouring pixels, *e.g.* Figure 6 (d). Figure 5 shows the workflow of the described process. The peripheral boundary ($PB$) which separates the breast interior from exterior ($BPA$) is extracted to facilitate intensity balancing, see Section 2.2.4 for details. To extract the $PB$, edge detection is firstly applied on the $BPA$. Then, for each image row, only the pixels closest to the breast interior are kept to form the boundary, *e.g.* Figure 6 (e).

#### 2.2.2. Intensity Ratio Propagation

Once the optimal $BPA$ is obtained, tissue appearance in the $BPA$ is improved by multiplying the original pixel greylevel value with a local intensity ratio calculated as a correction factor. A distance map (*e.g.* Figure 7 (c)) is firstly generated by

[2] Region filling is a form of mathematical morphology operator, which uses dilation as the basis, combined with logical operators.

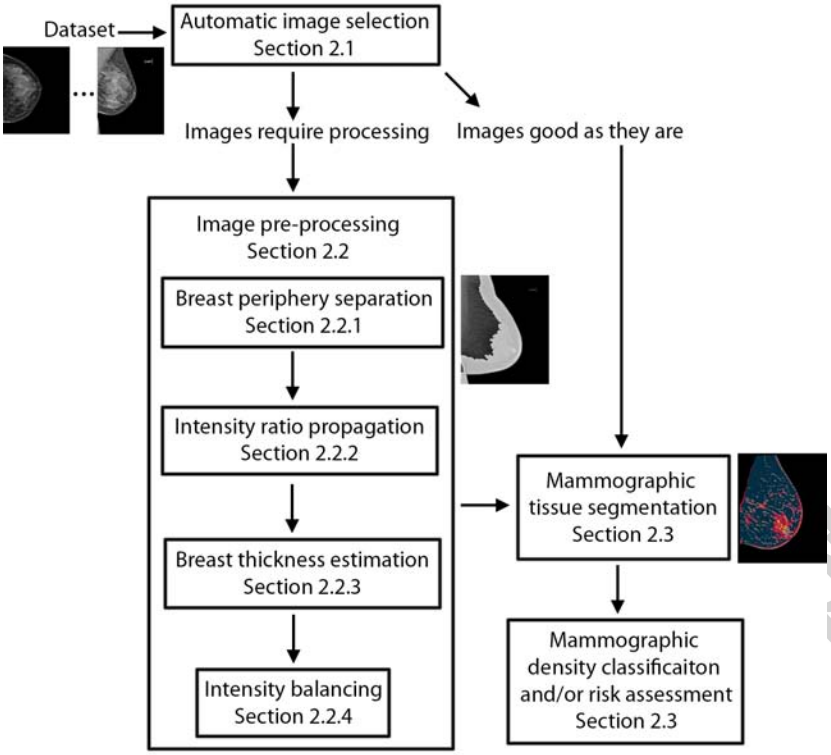
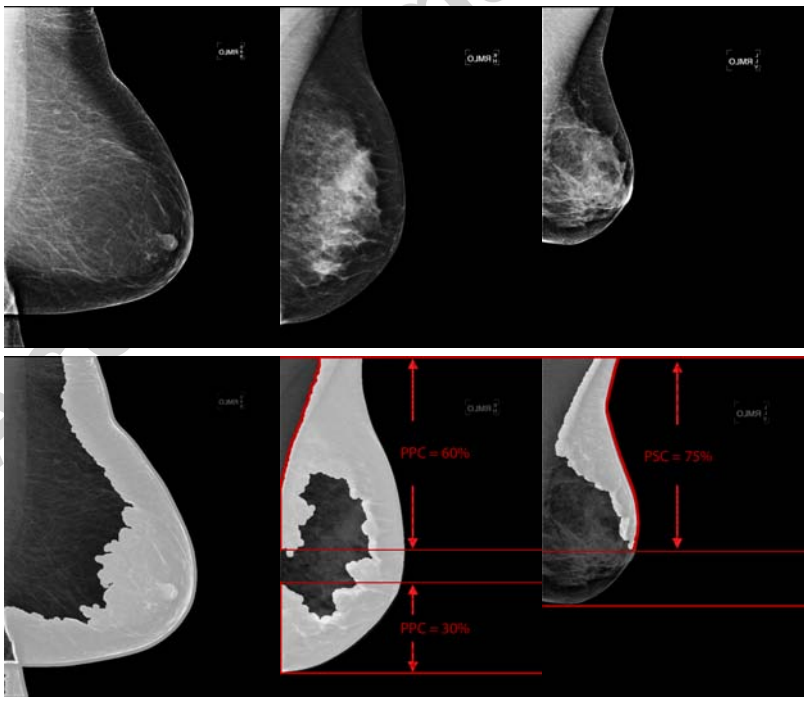



Figure 3: Illustration of overall workflow.




(a) (b) (c)

Figure 4: The top row shows the original images, the bottom row (a) shows superimposed white peripheral areas after Otsu binary segmentation, (b) shows 90% (60% + 30%) pectoral coverage of the image rows, and (c) shows 75% skinline coverage of the image rows.

| ID | Attributes | Minimum | Maximum | Mean | Standard deviation |
|---|---|---|---|---|---|
| A0 | patient's age | 42 | 91 | 60 | 10 |
| A1 | organ dose (dGy) | 0.008 | 0.0476 | 0.0184 | 0.0064 |
| A2 | entrance dose (mGy) | 3.2 | 31.2 | 10.7 | 4.4 |
| A3 | exposure (mAs) | 34 | 272 | 87.5 | 36.2 |
| A4 | relative X-ray exposure (mAs) | 279 | 580 | 419.8 | 61.6 |
| A5 | compression force (Newtons) | 44.5 | 249.1 | 103.5 | 34.5 |
| A6 | body part thickness (mm) | 29 | 104 | 61.7 | 14.8 |
| A7 | kVp | 24 | 35 | 30.4 | 2.3 |
| A8 | percent peripheral area % | 0.0 | 0.99 | 0.33 | 0.20 |
| A9 | percent skinline coverage % | 0.0 | 1.00 | 0.21 | 0.28 |
| A10 | percent pectoral coverage % | 0.0 | 0.96 | 0.23 | 0.22 |

Table 1: Statistics calculated for all the calibrated parameters and three additional image based parameters.

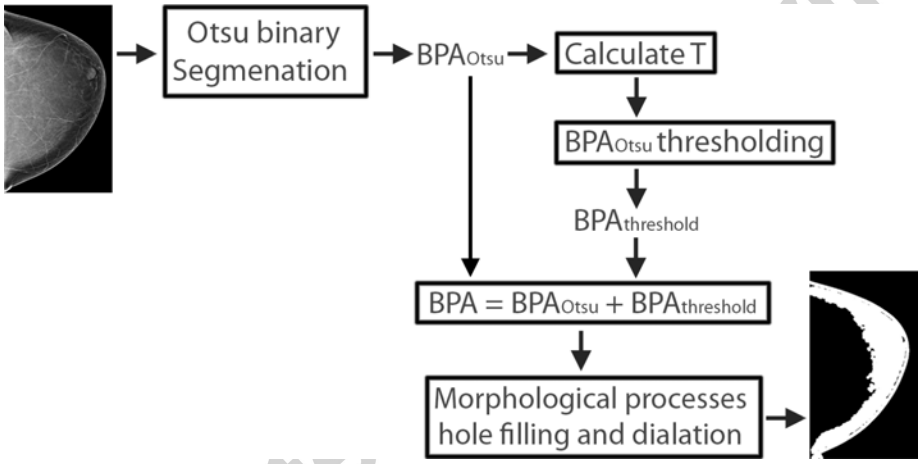



Figure 5: The workflow illustrates the BPA generation process.

(a) (b) (c) (d) (e)

Figure 6: From left to right showing; (a) the original image, (b) initial Otsu segmentation, (c) improved segmentation, (d) final breast peripheral area, and (e) peripheral boundary.

ACCEPTED MANUSCRIPT

calculating the shortest distance from each pixel $Img(x, y)$ to the skinline. Within the *BPA* (*e.g.* white area in Figure 7 (b)), the pixel corrections start from the pixels closest to the breast interior with the greatest distances to the skinline. For each pixel $P(x, y)$ with the distance to the skinline $D(x, y)$, the greylevel value is altered ($P'(x, y)$) by multiplying a propagation ratio ($pr$) calculated within an empirically defined $17 \times 17$ neighbourhood for efficiency and robustness as:

$$
\begin{aligned}
I_{avgP_2} &= \frac{\sum_{i=0}^{N} P_i(x,y)}{N}, \forall P_i(x,y) = D_i(x,y) + 2, \\
I_{avgP_1} &= \frac{\sum_{j=0}^{M} P_j(x,y)}{M}, \forall P_j(x,y) = D_j(x,y) + 1, \\
pr &= \frac{I_{avgP_2}}{I_{avgP_1}}, \\
P'(x,y) &= pr \times P(x,y),
\end{aligned}
$$

where $D(x, y) + 1$ and $D(x, y) + 2$ are pixel distances to the skinline 1 and 2 steps further away from $P(x, y)$. Figure 7 (d) shows an example image after intensity ratio propagation. The tissue structures in the *BPA* are noticeably improved when compared with the original image shown in Figure 7 (a).

### 2.2.3. Breast Thickness Estimation

The X-ray penetration strength has a direct correlation with breast thickness. Other physical properties (*e.g.* dosage, filter, and anode), unknown combination factors in the X-ray beam spectrum, and breast tissue composition may also affect mammographic appearance. In this work, these elements are encompassed in a 'black box' approach, and a non-linear relationship is assumed between tissue thickness and log-exposure (Beer's law of attenuation [16]).

To compensate the intensity variation due to breast thickness and tissue composition differences, a pair of CC and MLO is required to approximate the breast shape and estimate the relative breast thickness ratios. For example, for a CC view, the relative breast thickness ratio ($r$) can be estimated based on the projected physical contour of the compressed breast as seen on the MLO view. The skinline is extracted from the MLO view and split in two at the furthest pixel (at/near the nipple) to the chest wall to form the upper and lower skinlines (*e.g.* the blue and green lines in Figure 8 (b)). A chain code is generated for each skinline giving a sequence of pixels from start to end where the blue line meets the green line. For each pixel $P_i(x, y)$ in the top skinline, a corresponding pixel $P_j(x, y)$ is sought in the lower skinline, to form a parallel line (*pLine*) to the chest wall by linking the two pixels to form the longest line (*e.g.* red line in Figure 8 (c)). The slope ($m$) of the longest line is calculated as:

$$m = \frac{y_i - y_j}{x_i - x_j}.$$

The pixel linking process is repeated using the calculated $m$ for all the pixels in the chain code for the top skinline, resulting in a series of parallel lines (*e.g.* Figure 8 (c)-(e), all parallel to the longest line. For the CC view, the ratio $r$ at a given point $P$ (*e.g.* 'A' in Figure 8 (a)) is calculated based on the reference boundary pixel $P_{ref}$ (*e.g.* 'B' in Figure 8 (a)) as:

$$r = \frac{pLine(P)}{pLine(P_{ref})}.$$

Both pixels 'A' and 'B' in Figure 8 (a) are on the thickest projected section on the CC view (*e.g.* on the blue line in Figure 8 (a)). For the remaining pixels on the thickest projected section, the ratios are assigned in the same way. The calculated ratios are propagated to pixels with the same distance to the skinline (*e.g.* pixels on the yellow lines in Figure 8 (a) have the same distance to the skinline).

### 2.2.4. Intensity Balancing

The breast thickness ratios ($R(x, y)$) are log normalised based on the assumption of a non-linear relationship made between tissue thickness and log-exposure. The dotted red (L1) line shown in Figure 9 is for compressed breast thickness, and the solid blue (L2) line in Figure 9 is for log normalised breast thickness. To compensate intensity distribution variation based on $R(x, y)$, a global thickness reference ($R_{ref}$) is required and is the basis for all the corrections. $R_{ref}$ (*e.g.* intersection point between the solid blue (L2) and vertical cyan (L4) line in Figure 9) is defined as the mean thickness for all peripheral boundary pixels (*e.g.* cyan pixels in Figure 7 (b)). Once $R_{ref}$ is calculated, for each pixel $P'(x, y)$ within the *BPA* the greylevel value is altered as:

$$
\begin{aligned}
RP_{ref} &= \frac{R_{ref} - R_{min}}{R_{max} - R_{min}}, \\
RP(x,y) &= \frac{R(x,y) - R_{min}}{R_{max} - R_{min}}, \\
P''(x,y) &= P'(x,y)(1 + (RP_{ref} - RP(x,y))),
\end{aligned}
$$

where $RP_{ref}$ is the relative proportion of $R_{ref}$ to the overall breast thickness, $RP(x, y)$ is the relative proportion of breast thickness at $P'(x, y)$ to the overall breast thickness, and $P''(x, y)$ is the corrected pixel greylevel value. A higher or lower greylevel value is assigned if the relative breast thickness proportion is less or greater than $RP_{ref}$, respectively. Therefore, after correction, lower greylevel values within *BPA* increase (left to the vertical cyan (L4) line in Figure 9) and higher greylevel values outside *BPA* decrease (right to the vertical cyan (L4) line in Figure 9). The solid blue (L2) line (see Figure 9) move toward the 'balanced' breast thickness (*e.g.* the green (L3) line in Figure 9) to achieve intensity balancing with smooth intensity continuity. Figure 7 (a), (d), and (e) show an example image, the result after intensity ratio propagation, and the final result after intensity balancing, respectively. When comparing Figure 7 (d) with (e), (e) shows less 'over exposed' intensity in the centre of the breast. It should be noted that the example shown in Figure 7 is one of the most challenging images in the dataset, as it suffers from multiple issues; less visibility in the BPA and intensity imbalance across the image.

## 2.3. Mammographic Segmentation and Classification

All the processed mammographic images were segmented, followed by Tabár risk and BI-RADS density classification.

Accepted manuscript

ACCEPTED MANUSCRIPT

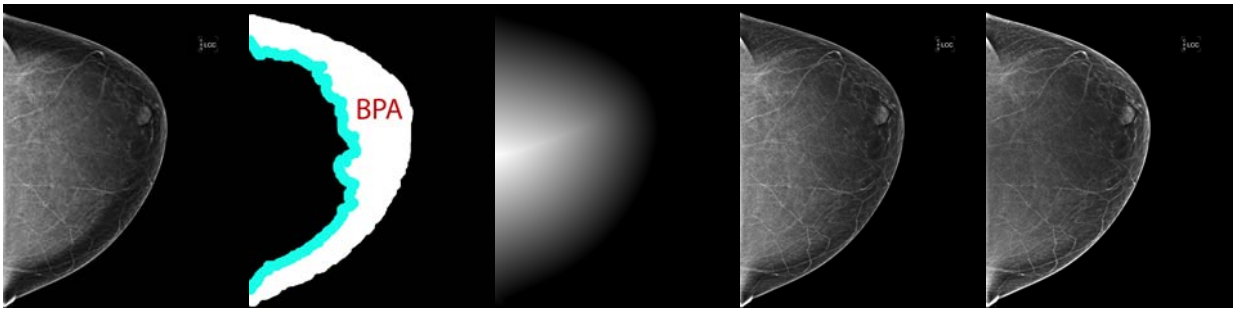



Figure 7: From left to right showing; (a) original image, (b) breast interior, peripheral boundary (cyan), and BPA (white), (c) distance map, (d) after intensity propagation, and (e) after intensity balancing.

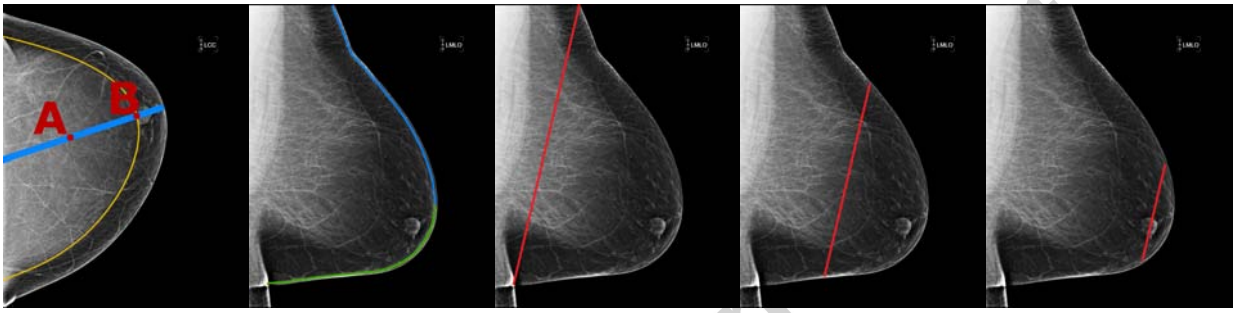



Figure 8: From left to right showing; (a) CC view, (b) paired MLO view, (c)-(e) parallel lines proximal to pectoral muscle, breast centre, and nipple.

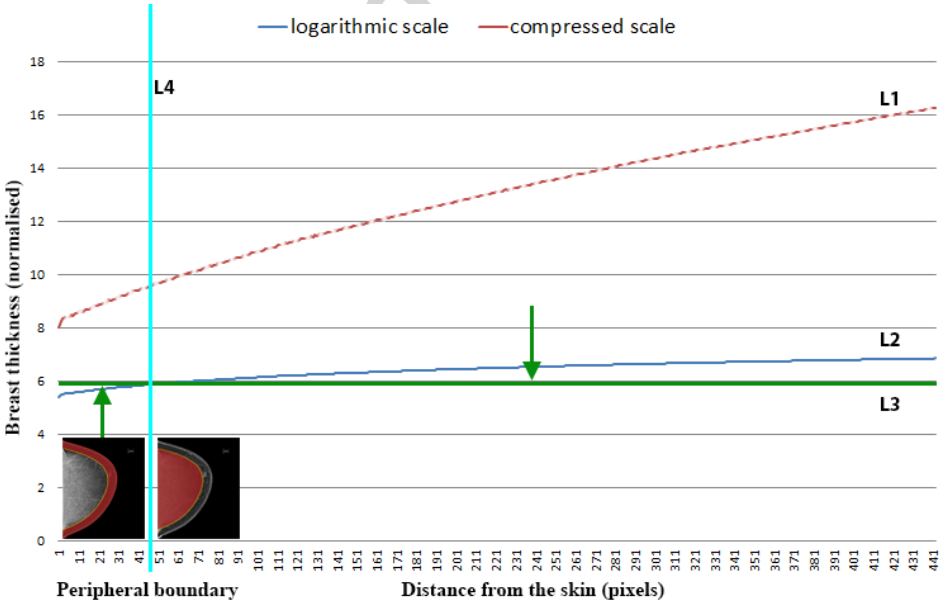



Figure 9: Approximated breast thickness (x-axis) when its compressed (dotted red L1), and log normalised (solid blue L2). The vertical cyan line (L4) indicates the boundary between BPA (dark red area in the left thumbnail) and breast interior (dark red area in the right thumbnail), based on the distances from the skinline (y-axis). The green (L3) line is the 'idea' balanced (even) breast thickness.

ACCEPTED MANUSCRIPT

Previously investigated greylevel histogram [17] and geometric moments [18] based features were extracted from four sets of mammographic patches, containing samples of (139) nodular, (224) linear, (87) homogeneous, and (89) radiolucent tissue. A total of 23 features were used which are expected to contain not only texture primitives but also geometric information. A feature and classifier selection process was incorporated using a collection of attribute selection algorithms and classifiers available in Weka (see full list in [15]). In particular, a set of neighbourhoods (i.e. {7, 17, 27, 37}) covering small to large anatomical structures were used in the feature extraction. The derived feature vectors were subjected to all available filtering methods in Weka for feature selection and dimensionality reduction, in order to select the most discriminative subset of the features. All available classifiers in Weka were used to perform (10-fold) cross-validation based evaluation over the selected features based on half a million randomly selected pixels from the 539 patches. Emperical testing indicated that the highest classification results can be achieved using the random committee algorithm with an average accuracy ~79% (based on 5 iterations), which was used in conjunction with the selected features for Tabár tissue modelling.

A model driven pixel based segmentation was performed using the selected classifier. The derived breast tissue composition was compared with Tabár and BI-RADS schemes (empirical clinical models) in order to find the closest matches in the Euclidean space using a nearest neighbour classifier, as a means of mammographic risk/density classification.

Tabár's scheme was used to facilitate the evaluation due to its models being quantitatively defined. The closely related BI-RADS scheme is chosen as a means of performing comparisons between different risk/density assessment schemes. It should be noted that Tabár/BI-RADS based four-class tissue segmentation might not directly translate to other breast density measures; *e.g.* Cumulus [19], Volpara™ [20], and Quantra™ [21], see recent work on the reliability of automated breast density measurements [22].

## 3. Results and Discussion

This section presents evaluation results, including: 1) the accuracies in selecting problematic images for pre-processing prior to analysis; 2) mammographic segmentation when using the processed images; 3) and subsequent risk/density classification. In addition, clinical evaluations were conducted which visually assessed the quality of the processed and segmented mammographic images.

### *3.1. Automatic Image Selection*

Automatic image selection was able to correctly identify 93% of images requiring pre-processing. Most misclassified images have appearances somewhere between 'good as it is' (GAII) and 'requiring per-processing' (RPP). Table 2 shows the classification confusion matrix, the derived Kappa statistic ($\kappa$ = 0.86) indicates an almost perfect agreement. Figure 10 shows plots of all feature pairs for the attributes shown in Table 1. Blue and red dots represent images as GAII and RPP, respectively. Based on the separation between these two groups, attributes such as compression force (A6), kVp (A7), PPC (A9), and PSC (A10) may provide more discriminating power in identifying problematic images. On the other hand, patient's age (A0), PPA (A8), relative X-ray exposure (A4), and compression force (A5) seems to be less robust in separating the two groups. Using kVp can lead to a better separation in the feature space, this is in line with the observation made in the related studies [23, 24]. It was expected that compression force may have a more direct impact in image quality, but this is not clearly demonstrated in the scatter plots in Figure 10.

| | GAII | RPP |
|---|---|---|
| GAII | 169 | 13 |
| RPP | 13 | 165 |

Table 2: Classification confusion matrix; 'good as it is' (GAII) (♯ images = 182) and 'requiring per-processing' (RPP) (♯ images = 178) images.

| | category A | category B | category C |
|---|---|---|---|
| ♯ images | 25 | 88 | 65 |
| Accuracies | 14% | 49% | 37% |

Table 3: Image quality evaluation for the processed images when compared with the original images before pre-processing. Note that there are 178 (49%) images requiring processing.

### *3.2. Image Pre-processing*

Each processed image was rated as 'interfere with image interpretation' (category A), 'the same interpretation' (category B), or 'improvement in image interpretation' (category C). It should be noted that category B can mean that the processed images have no apparent visual concerns/improvements in image interpretation. It should be noted that there is no image quality evaluation standard exist in a clinical environment. The categories were defined based on the difficulty of mammographic image interpretation by the standard of consultant radiologists. Each image was rated using visual assessment by the radiologists. Table 3 shows the rating results when compared with the original images before pre-processing, and indicates a relatively small negative impact (i.e. 14% of images in category A) after pre-processing. In most cases (i.e. 86% of images in categories B and C) the processed images can be interpreted the same as using the original ones or have improved interpretation. Figure 11 (a)-(c) shows example processed images of categories A-C, respectively. Intensity ratio propagation was able to improve the BPA visibility for most of the cases, see Figure 11 (a) for example. It performed less satisfactorily for images exhibiting extensive BPA, in some cases the BPA can occupy almost half of the breast and the correction may not work consistently, for example Figure 11 (b) shows better corrected BPA in the top half of the breast but not as well for the bottom half of it. Intensity balancing may be less robust in correcting greylevel values of areas near the axillary tail, when compared with that in the central breast region, see Figure 11 (b) for example. Figure 11 (c) shows an example where enhancement may not be nec-

Accepted manuscript

ACCEPTED MANUSCRIPT

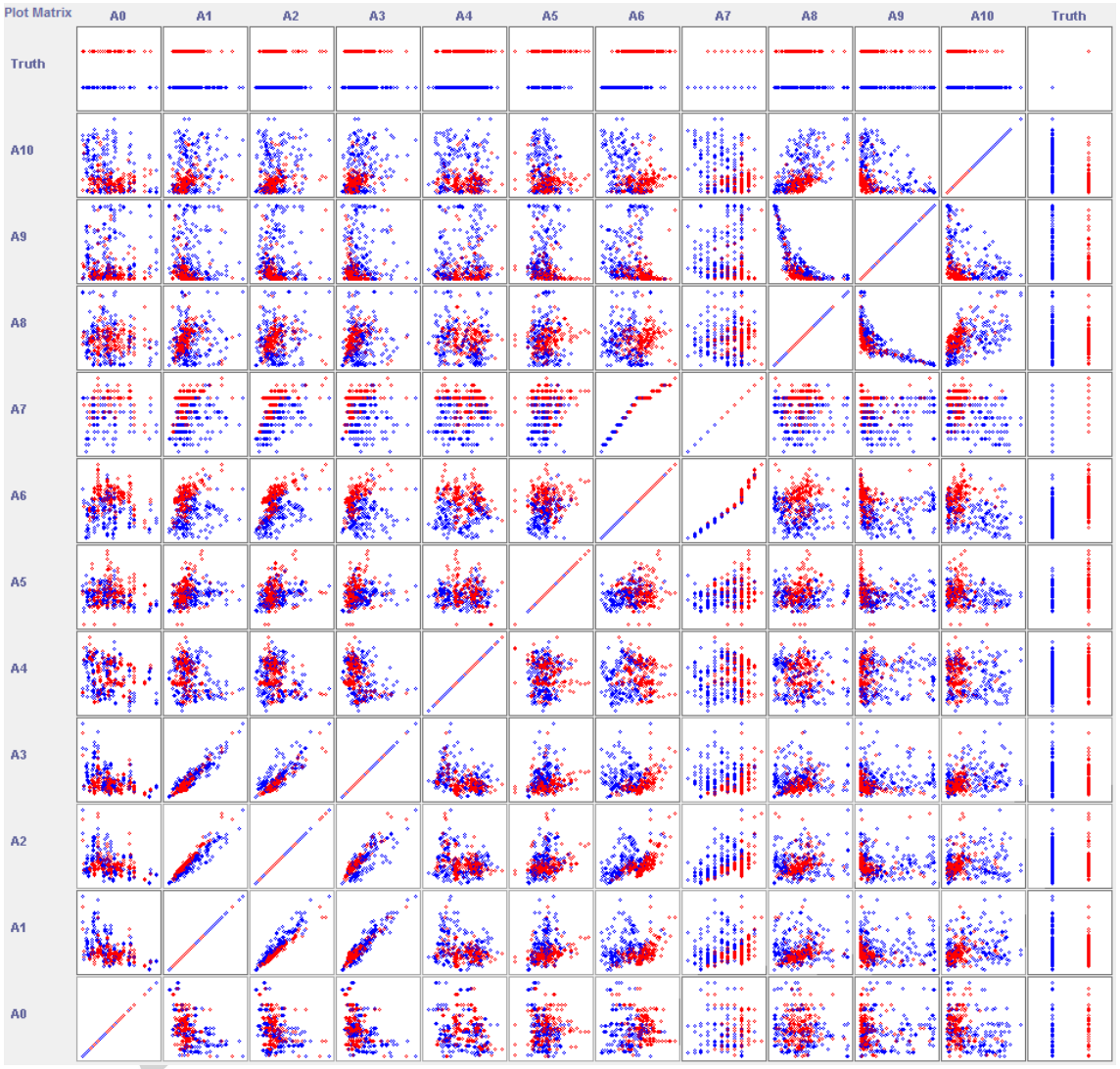
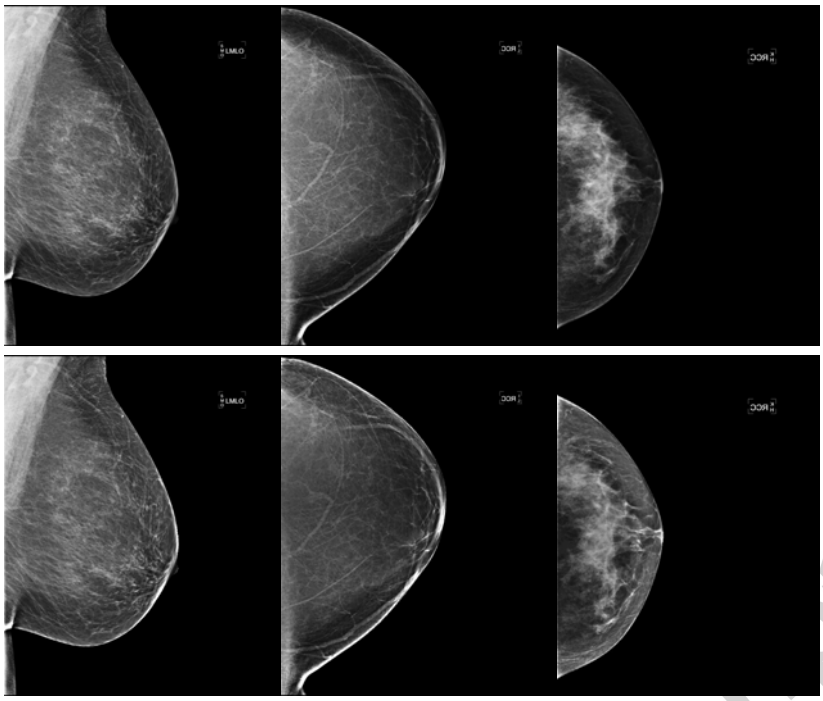



Figure 10: Attributes plot matrix. A0 to A10 (see Table 1) are patient's age, organ dose (dGy), entrance dose (mGy), exposure (mAs), relative X-ray exposure (mAs), compression force (Newtons), body part thickness (mm), kVp, percent peripheral area (PPA) %, percent skinline coverage (PSC) %, percent pectoral coverage (PPC) %, respectively. Blue and red dots represent images as 'good as it is' (GAII) and 'requiring per-processing' (RPP), respectively.

ACCEPTED MANUSCRIPT



(a) (b) (c)

Figure 11: Example processed images; the top and bottom rows show the original and processed images, respectively; from left to right the correction process has a positive, neutral, and negative effect. Images from left to right were rated $B_1/T_I$, $B_1/T_{II}$, and $B_2/T_I$, respectively. Note that most of the problematic images belong to $B_1$, $B_2$, $T_I$, and $T_{II/III}$.

essary, in this case the process caused distortions to the original image's appearance. The correction process has the ability to improve mammographic appearances but one potential issue can be seen; the process may distort abnormalities (*e.g.* spiculated mass and circumscribed lesion) and contrast of underlying anatomical structures. This can affect screening interpretation but may not be an issue for certain image processing tasks (*e.g.* segmentation). Intensity ratio propagation can introduce ringing artefacts in the BPA, however, this is less noticeable with the used parameter configurations. It should be noted that when compared to early development in [23, 24], the current implementation is substantially different; it has fewer stages and is more robust in ensuring the intensity continuation during the correction.

### *3.3. Segmentation*

Figure 12 shows example mammographic segmentation. When using the processed images, results show that there are less missegmented pixels near the chest wall (thicker part of breast), at the same time more linear structures (*e.g.* blood vessels) were segmented in the peripheral area (*e.g.* Figure 12 (4)-(a-d)). In some cases missegmentation can be seen near the skinline (*e.g.* Figure 12 (a)-(b)) due to the artefacts (brighter skinline) created after intensity propagation. Overall, when using the processed images, results show more detailed segmentation for nodular and homogeneous tissue, and missegmented linear density (*e.g.* Figure 12 (4)-(b)). Note that linear structures can be hard to segment because this type of tissue is often subtly embedded in other types of tissue. Tabár's mammographic building blocks; radiolucent, linear, nodular, and homogeneous tissue can be loosely mapped to BI-RADS density categories as fatty, semi-fatty, dense, and semi-dense, respectively. All the segmented images were clinically rated based on the correctness of the segmented density categories corresponding to the BI-RADS density categories. The clinical evaluation was conducted based only on the BI-RADS scheme because this scheme has been widely used in the US and some European countries, whilst the Tabár scheme has not (yet) been adopted in a clinical environment. Therefore, clinical evaluation with respect to the Tabár scheme is not included in the current study. During the evaluation, three ratings were considered; 'Unacceptable/Poor' (U/P), 'Acceptable' (A), and 'Good/Excellent' (G/E). Tables 4 (a) and (b) show the results for the entire dataset (♯ images = 360) before and after pre-processing. Table 4 (a) and (b) indicate that there is a 27% improvement in the A and G/E categories. Table 4 (c)-(f) show rating results with respect to the BI-RADS density categorises after pre-processing. We observed a 99% correct fatty segmentation within A (8%) and G/E (91%) categories in Table 4 (c). A good fatty segmentation is expected as this tissue type is relatively easy to be identified. Table 4 (d) shows a 60% correct semi-fatty segmentation within the A (37%) and G/E (23%) categories. The decrease in segmentation accuracies for this type of tissue can be related to the fact that it is difficult to correctly identify tissue during its transitional stage (e.g. during hormone replacement therapy and tissue change due to natural ageing). Table 4 (e) shows a 73% correct semi-dense segmentation within the A (20%) and G/E (53%) categories. Table 4 (f) shows a 91% correct dense

Accepted manuscript

| | U/P | A | G/E |
|---|---|---|---|
| $B_1$ | 133 (78%) | 22 (13%) | 15 (9%) |
| $B_2$ | 37 (30%) | 37 (30%) | 49 (40%) |
| $B_2$ | 11 (25%) | 14 (33%) | 18 (42%) |
| $B_3$ | 13 (54%) | 10 (42%) | 1 (4%) |
| Total | 194 (54%) | 83 (23%) | 83 (23%) |

(a) before pre-processing

| | U/P | A | G/E |
|---|---|---|---|
| $B_1$ | 78 (46%) | 59 (35%) | 33 (19%) |
| $B_2$ | 17 (14%) | 43 (35%) | 63 (51%) |
| $B_3$ | 3 (7%) | 8 (19%) | 32 (74%) |
| $B_4$ | 0 (0%) | 13 (54%) | 11 (46%) |
| Total | 98 (27%) | 123 (34%) | 139 (39%) |

(b) after pre-processing

| | U/P | A | G/E |
|---|---|---|---|
| $B_1$ | 4 (2%) | 18 (11%) | 148 (87%) |
| $B_2$ | 1 (1%) | 9 (7%) | 113 (92%) |
| $B_3$ | 0 (0%) | 0 (0%) | 43 (100%) |
| $B_4$ | 0 (0%) | 0 (0%) | 24 (100%) |
| Total | 5 (1%) | 27 (8%) | 328 (91%) |

(c) fatty tissue

| | U/P | A | G/E |
|---|---|---|---|
| $B_1$ | 99 (58%) | 47 (28%) | 24 (14%) |
| $B_2$ | 33 (27%) | 55 (45%) | 35 (28%) |
| $B_3$ | 6 (14%) | 18 (42%) | 19 (44%) |
| $B_4$ | 7 (29%) | 14 (58%) | 3 (13%) |
| Total | 145 (40%) | 134 (37%) | 81 (23%) |

(d) semi-fatty

| | U/P | A | G/E |
|---|---|---|---|
| $B_1$ | 85 (50%) | 40 (24%) | 45 (26%) |
| $B_2$ | 13 (10%) | 25 (20%) | 85 (70%) |
| $B_3$ | 0 (0%) | 4 (10%) | 39 (90%) |
| $B_4$ | 0 (0%) | 1 (4%) | 23 (96%) |
| Total | 98 (27%) | 70 (20%) | 192 (53%) |

(e) semi-dense tissue

| | U/P | A | G/E |
|---|---|---|---|
| $B_1$ | 11 (7%) | 79 (46%) | 80 (47%) |
| $B_2$ | 12 (10%) | 61 (50%) | 50 (40%) |
| $B_3$ | 7 (16%) | 16 (37%) | 20 (47%) |
| $B_4$ | 2 (8%) | 20 (83%) | 2 (9%) |
| Total | 32 (9%) | 176 (49%) | 152 (42%) |

(f) dense tissue

Table 4: Clinical ratings for the quality of the segmented mammographic images using the full dataset (♯ images = 360). 'U/P', 'A', and 'G/E' denote 'Unacceptable/Poor', 'Acceptable', and 'Good/Excellent', respectively. Note that (c)-(f) results are after pre-processing, and represent fatty, semi-fatty, semi-dense, and dense tissue, respectively.

| | U/P | A | G/E |
|---|---|---|---|
| $B_1$ | 95 (76%) | 18 (14%) | 12 (10%) |
| $B_2$ | 21 (41%) | 15 (29%) | 15 (30%) |
| $B_2$ | 2 (100%) | 0 (0%) | 0 (0%) |
| $B_3$ | 0 (0%) | 0 (0%) | 0 (0%) |
| Total | 118 (66%) | 33 (19%) | 27 (15%) |

(a) before pre-processing

| | U/P | A | G/E |
|---|---|---|---|
| $B_1$ | 65 (52%) | 46 (37%) | 14 (11%) |
| $B_2$ | 9 (18%) | 25 (49%) | 17 (33%) |
| $B_3$ | 0 (0%) | 0 (0%) | 2 (100%) |
| $B_4$ | 0 (0%) | 0 (0%) | 0 (0%) |
| Total | 74 (41%) | 71 (40%) | 33 (19%) |

(b) after pre-processing

| | U/P | A | G/E |
|---|---|---|---|
| $B_1$ | 2 (2%) | 12 (10%) | 111 (88%) |
| $B_2$ | 1 (2%) | 5 (10%) | 45 (88%) |
| $B_3$ | 0 (0%) | 0 (0%) | 2 (100%) |
| $B_4$ | 0 (0%) | 0 (0%) | 0 (0%) |
| Total | 3 (2%) | 17 (10%) | 158 (88%) |

(c) fatty tissue

| | U/P | A | G/E |
|---|---|---|---|
| $B_1$ | 90 (72%) | 24 (19%) | 11 (9%) |
| $B_2$ | 24 (47%) | 20 (39%) | 7 (14%) |
| $B_3$ | 0 (0%) | 2 (100%) | 0 (0%) |
| $B_4$ | 0 (0%) | 0 (0%) | 0 (0%) |
| Total | 114 (64%) | 46 (26%) | 18 (10%) |

(d) semi-fatty

| | U/P | A | G/E |
|---|---|---|---|
| $B_1$ | 68 (54%) | 32 (26%) | 25 (20%) |
| $B_2$ | 7 (14%) | 12 (24%) | 32 (62%) |
| $B_3$ | 0 (0%) | 0 (0%) | 2 (100%) |
| $B_4$ | 0 (0%) | 0 (0%) | 0 (0%) |
| Total | 75 (42%) | 44 (25%) | 59 (33%) |

(e) semi-dense tissue

| | U/P | A | G/E |
|---|---|---|---|
| $B_1$ | 5 (4%) | 58 (46%) | 62 (50%) |
| $B_2$ | 5 (10%) | 20 (40%) | 26 (50%) |
| $B_3$ | 0 (0%) | 0 (0%) | 2 (100%) |
| $B_4$ | 0 (0%) | 0 (0%) | 0 (0%) |
| Total | 10 (6%) | 78 (44%) | 90 (50%) |

(f) dense tissue

Table 5: Clinical ratings for the quality of the segmented mammographic images using images 'requiring pre-processing' (♯ images = 178). 'U/P', 'A', and 'G/E' denote 'Unacceptable/Poor', 'Acceptable', and 'Good/Excellent', respectively. Note that (c)-(f) results are after pre-processing, and represent fatty, semi-fatty, semi-dense, and dense tissue, respectively.

Accepted manuscript

ACCEPTED MANUSCRIPT

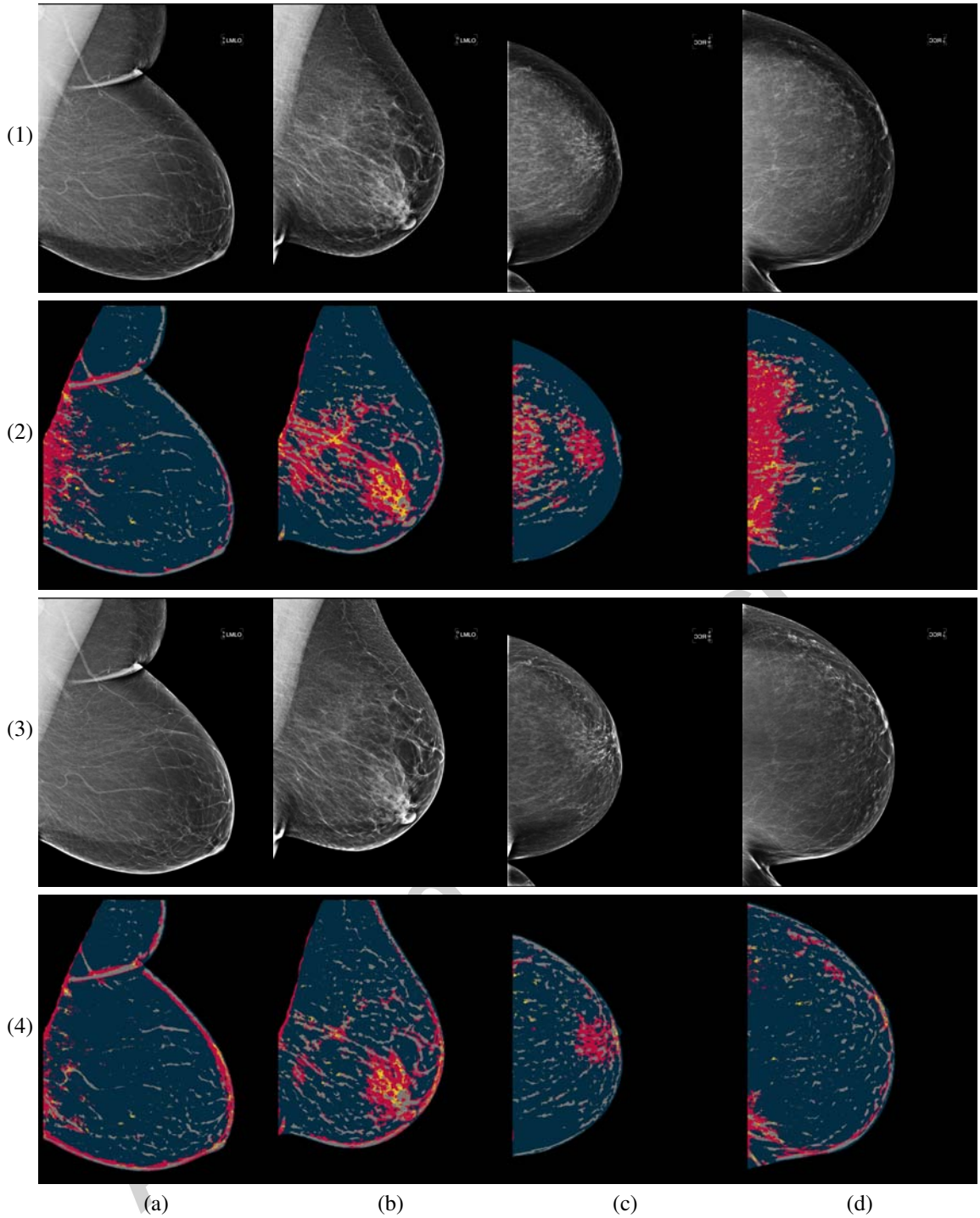



Figure 12: Example mammographic segmentation. From top to bottom showing the original images (1), its segmentation (2), the processed images (3), and the corresponding segmentation (4). Note that images (a)-(d) were rated $B_1/T_{III}$, $B_1/T_{III}$, $B_1/T_I$, and $B_1/T_{III}$, respectively.

| | $T_I$ | $T_{II/III}$ | $T_{IV}$ | $T_V$ |
|---|---|---|---|---|
| $T_I$ | 106 | 18 | 13 | 0 |
| $T_{II/III}$ | 22 | 131 | 4 | 1 |
| $T_{IV}$ | 18 | 4 | 32 | 0 |
| $T_V$ | 2 | 2 | 7 | 0 |

(MLO/CC)

| | $T_I$ | $T_{II/III}$ | $T_{IV}$ | $T_V$ |
|---|---|---|---|---|
| $T_I$ | 45 | 15 | 7 | 0 |
| $T_{II/III}$ | 8 | 67 | 3 | 0 |
| $T_{IV}$ | 3 | 2 | 23 | 2 |
| $T_V$ | 0 | 1 | 1 | 3 |

(CC)

| | $T_I$ | $T_{II/III}$ | $T_{IV}$ | $T_V$ |
|---|---|---|---|---|
| $T_I$ | 57 | 10 | 3 | 0 |
| $T_{II/III}$ | 9 | 71 | 0 | 0 |
| $T_{IV}$ | 12 | 3 | 7 | 2 |
| $T_V$ | 2 | 2 | 2 | 0 |

(MLO)

Table 6: Risk classification confusion matrices using the Tabár scheme. From left to right; images and the Kappa statistics are MLO/CC ($\kappa$ = 0.60), CC ($\kappa$ = 0.64), and MLO ($\kappa$ = 0.59) views, respectively.

ACCEPTED MANUSCRIPT

| | $B_1$ | $B_2$ | $B_3$ | $B_4$ |
|---|---|---|---|---|
| $B_1$ | 149 | 16 | 2 | 3 |
| $B_2$ | 23 | 90 | 10 | 0 |
| $B_3$ | 4 | 23 | 13 | 3 |
| $B_4$ | 3 | 3 | 1 | 17 |

(MLO/CC)

| | $B_1$ | $B_2$ | $B_3$ | $B_4$ |
|---|---|---|---|---|
| $B_1$ | 72 | 12 | 0 | 1 |
| $B_2$ | 13 | 43 | 5 | 1 |
| $B_3$ | 1 | 8 | 10 | 2 |
| $B_4$ | 0 | 4 | 1 | 7 |

(CC)

| | $B_1$ | $B_2$ | $B_3$ | $B_4$ |
|---|---|---|---|---|
| $B_1$ | 67 | 15 | 2 | 1 |
| $B_2$ | 8 | 43 | 8 | 2 |
| $B_3$ | 1 | 10 | 11 | 0 |
| $B_4$ | 2 | 2 | 1 | 7 |

(MLO)

Table 7: Density classification confusion matrices using the BI-RADS scheme. From left to right; images and the Kappa statistics are MLO/CC ($\kappa$ = 0.60), CC ($\kappa$ = 0.58), and MLO ($\kappa$ = 0.55) views, respectively.

segmentation within the A (49%) and G/E (42%) categories. The break down results show overall satisfactory ratings for fatty and dense tissue segmentation. However, the U/P results in Tables 4 (e) and (f) indicate that the developed approach is robust to identify semi-dense tissue, but a decision line between semi-dense and dense tissue can be hard to determine as there are more 'acceptable' ratings for dense tissue which indicate the radiologists may have a neutral feeling about the segmentation. Tables 5 (a) and (b) show rating results for the images automatically identified as 'requiring pre-processing' (♯ images = 178) before and after pre-processing, respectively. Results are similar when compared with those derived from the full dataset, however, it should be noted that automatically selected images 'requiring pre-processing' mainly consisting of $B_1$ and $B_2$. Clinical feedback with respect to the evaluation based on the BI-RADS scheme indicated that problematic cases show difficulty in differentiating between semi-fatty and semi-dense tissue in the lateral and medial aspects of the breast; this is reflected in poor U/P results in Table 4 (d) and (e). Semi-dense tissue at the lower back can be a little too prominent which was reflected in over-segmentation near sub dermal (skin) and posterior fatty breast areas. In some cases there was a lack of semi-fatty tissue near the lateral back, and segmentation showed not enough dense 'islands' due to semi-fatty inhomogeneous regions. Medial vessel (linear tissue) caused 'spilled' over density and some fine fibrous structures may not be segmented posteriorly.

### *3.4. Risk/Density Classification*

Tables 6 and 7 show mammographic risk classification results when using the Tabár and BI-RADS schemes. The classification comparison before and after the pre-processing can be found in Table 8. The total classification accuracies were 75% (CC/MLO), 77% (CC), and 75% (MLO) when the results were evaluated using the Tabár's scheme. There is an average 7% increase when compared with results obtained for the original images. The total classification accuracies were 75% (MLO/CC), 73% (CC), and 71% (MLO) when the results were evaluated using the BI-RADS scheme. There is an average 5% increase when compared with results obtained for the original images. The results show lower accuracies for the MLO cases after pre-processing. This may indicate that the developed pre-processing can better estimate the breast thickness due to the single concave breast outline in the CC view, while the estimation is less precise over the more complicated concave and convex breast outline in the MLO view. During breast thickness estimation (see Section 2.2.3), a series of lines are generated and parallel to the chest wall. The longest parallel line is used as reference for the generation of the rest of the parallel lines. In almost all CC view cases, the longest parallel line is vertical due to the CC view orientation. However, it can be problematic for MLO view cases, as the generated reference parallel lines are not always accurate due to slightly 'angled' (*e.g.* leaning forward) breast with varying degrees. This may produce incorrect breast thickness estimation for the MLO view cases, leading to negative effect in segmentation and decrease in classification accuracy. As future work, a dedicated method can be introduced as part of the process to correctly identify the chest wall and check the parallel lines' alignment. Mammographic images in high risk/density categories seem to have more misclassification (percentage wise), which may relate to the intensity over balancing for structureless dense tissue (*e.g.* nodular and homogeneous). $T_{IV}$ is more likely to be classified into $T_I$. This may be due to the tissue distributions for these two patterns being more similar than others. Whilst with the BI-RADS scheme, dense tissue proportion increases with risk, therefore, most misclassification can be found in either a density category higher or lower (*e.g.* $B_2$ misclassified into $B_1$ or $B_3$). It should be noted that Tabár low risk categories (i.e. $T_I$ and $T_{II/III}$ ) are not directly correlated with BI-RADS low density categories (i.e. $B_1$ and $B_2$) [25]. Table 9 shows risk classification accuracies based on high and low risk categories. The results for the CC view are more accurate than the CC/MLO combination. This may be related to the more complicated concave and convex breast outline seen in the MLO view as discussed previously. The results are close to the results achieved (i.e. on average > 80%) using the state-of-the-art method [6]. It should be noted that different segmentation (*e.g.* fatty/dense segmentation) and classification (*e.g.* high/low risk classification) principle and datasets were used in [6].

## 4. Conclusions

The developed mammographic image pre-processing technique showed the ability to improve the contrast of tissue structures in uncompressed breast peripheral areas, and at the same time reduce intensity discrepancies across the mammograms. The novel automatic selection approach was able to better target images requiring pre-processing in a systematic way. A quantitative and qualitative evaluation was conducted to assess the usefulness of the developed method. When using the processed mammographic images for segmentation, there are more anatomically accurate and consistent results over the breast

Accepted manuscript

ACCEPTED MANUSCRIPT

| | MLO/CC | CC | MLO |
|---|---|---|---|
| Tabár | 69% | 68% | 79% |
| BI-RADS | 69% | 67% | 77% |

| | MLO/CC | CC | MLO |
|---|---|---|---|
| Tabár | 75% | 77% | 75% |
| BI-RADS | 75% | 73% | 71% |

Table 8: Tabár risk/BI-RADS density classification accuracies before (left) and after (right) pre-processing.

| | MLO/CC | CC | MLO |
|---|---|---|---|
| Tabár | 88% | 91% | 88% |
| BI-RADS | 87% | 89% | 84% |

Table 9: Tabár risk/BI-RADS density classification accuracies based on high and low categories. High risks; $T_{IV}$ and $T_V$. Low risks; $T_I$ and $T_{II/III}$. High densities; $B_3$ and $B_4$. Low densities; $B_1$ and $B_2$.

parenchyma, this in turn improved subsequent risk classification accuracies. There are significant positive relationships between the radiologists manual and automatic mammographic risk assessments for both Tabár (substantial agreement) and BI-RADS (moderate agreement) schemes. Clinical evaluation showed that pre-processing can have positive impacts on mammographic segmentation, however, the processed images are not ready to be used for interpretation. Further validation in a clinical environment is required in order to extend the usage for mammographic reading purposes; in addition, investigations into multi-vendor evaluation and density estimation comparison with other approaches should be considered. Utilising such an image pre-processing technique in a mammographic segmentation methodology can prove useful in quantifying change in relative proportion of breast tissue, aiding radiologists' estimation in mammographic risk/density categories, and providing risk-stratified screening for patients.

## References


[1] Office for National Statistics, Cancer Statistics Registrations, England (Series MB1), Cancer Research UK (44) (2013) 1.

[2] J. J. Fenton, S. H. Taplin, P. A. Carney, L. Abraham, E. A. Sickles, C. D'Orsi, E. A. Berns, G. Cutter, R. E. Hendrick, W. E. Barlow, J. G. Elmore, Influence of computer-aided detection on performance of screening mammography, New England Journal of Medicine 356 (14) (2007) 1399–1409.

[3] U. Bick, M. L. Giger, R. A. Schmidt, R. M. Nishikawa, K. Doi, Density correction of peripheral breast tissue on digital mammograms., RadioGraphics 16 (6) (1996) 1403–1411.

[4] J. W. Byng, J. P. Critten, M. J. Yaffe, Thickness-equalization processing for mammographic images, Radiology 203 (2) (1997) 564–568.

[5] P. R. Snoeren, N. Karssemeijer, Thickness correction of mammographic images by anisotropic filtering and interpolation of dense tissue, Processing of SPIE, Medical Imaging: Image Processing 5747 (2005) 1521–1527.

[6] M. Tortajada, A. Oliver, R. Martí, S. Ganau, L. Tortajada, M. Sentís, J. Freixenet, R. Zwiggelaar, Breast peripheral area correction in digital mammograms, Computers in Biology and Medicine 50 (2014) 32–40.

[7] P. R. Snoeren, N. Karssemeijer, Thickness correction of mammographic images by means of a global parameter model of the compressed breast, IEEE Transactions on Medical Imaging 23 (7) (2004) 799–806.

[8] A. P. Stefanoyiannis, L. Costaridou, S. Skiadopoulos, G. Panayiotakis, A digital equalisation technique improving visualisation of dense mammary gland and breast periphery in mammography, European Journal of Radiology 45 (2) (2003) 139–149.

[9] O. Alonzo-Proulx, J. G. Mainprize, N. J. Packard, J. M. Boone, A. Al-Mayah, K. Brock, M. J. Yaffe, Development of a peripheral thickness estimation method for volumetric breast density measurements in mammography using a 3d finite element breast model, Lecture Notes in Computer Science 6136 (2010) 467–473.

[10] M. G. J. Kallenberg, N. Karssemeijer, Compression paddle tilt correction in full-field digital mammograms, Physics in Medicine and Biology 57 (3) (2012) 703–715.

[11] C. E. Tromans, M. R. Cocker, S. M. Brady, Quantification and normalization of x- ray mammograms, Physics in Medicine and Biology 57 (20) (2012) 6519–6540.

[12] L. Tabár, T. Tot, P. B. Dean, Breast Cancer: The Art And Science Of Early Detection With Mamography: Perception, Interpretation, Histopatholigic Correlation, 1st Edition, Georg Thieme Verlag, 2004.

[13] American College of Radiology, Breast Imaging Reporting and Data System BI-RADS, 4th Edition, Reston, VA: American College of Radiology, 2004.

[14] M. Sezgin, B. Sankur, Survey over image thresholding techniques and quantitative performance evaluation, Journal of Electronic Imaging 13 (2004) 146–165.

[15] E. F. I. H. Witten, M. A. Hall, Data Mining: Practical machine learning tools and techniques, 3rd Edition, Morgan Kaufmann, San Francisco, 2011.

[16] Beer, Bestimmung der absorption des rothen lichts in farbigen flüssigkeiten, Annalen der Physik 162 (1852) 78–88.

[17] W. He, E. R. E. Denton, R. Zwiggelaar, Mammographic segmentation and risk classification using a novel binary model based Bayes classifier, Lecture Notes in Computer Science 7361 (2012) 40–47.

[18] W. He, E. R. E. Denton, K. Stafford, R. Zwiggelaar, Mammographic image segmentation and risk classification based on mammographic parenchymal patterns and geometric moments, Biomedical Signal Processing and Control 6 (3) (2011) 321–329.

[19] N. Boyd, J. Byng, R. Jong, E. Fishell, L. Little, A. Miller, G. Lockwood, D. Tritchler, M. J. Yaffe, Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian national breast screening study, Journal of the National Cancer Institute 8787 (1995) 670–675.

[20] R. Highnam, M. Brady, M. J. Yaffe, N. Karssemeijer, J. Harvey, Robust breast composition measurement - Volpara$^{TM}$, Lecture Notes in Computer Science 6136 (2010) 342–349.

[21] K. Hartman, R. Highnam, R. Warren, V. Jackson, Volumetric assessment of breast tissue composition from FFDM images, Lecture Notes in Computer Science 5116 (2008) 33–39.

[22] O. Alonzo-Proulx, G. E. Mawdsley, J. T. Patrie, M. J. Yaffe, J. A. Harvey, Reliability of automated breast density measurements, Radiology 275 (2) (2015) 366–376.

[23] W. He, M. Kibiro, A. Juette, E. R. E. Denton, P. Hogg, R. Zwiggelaar, A novel image enhancement methodology for full field digital mammography, Lecture Notes in Computer Science 8539 (2014) 650–657.

[24] W. He, E. R. E. Denton, R. Zwiggelaar, A novel breast image preprocessing for full field digital mammographic segmentation and risk classification, Medical Image Understanding and Analysis (2014) 40–47.

[25] W. He, M. Kibiro, A. Juette, E. R. E. Denton, R. Zwiggelaar, A revisit on correlation between Tabár and Birads based risk assessment schemes with full field digital mammography, Lecture Notes in Computer Science 8539 (2014) 327–333.


Accepted manuscript